AD-767 189

PSEUDO STATE MEASUREMENTS APPLED TO RECURSIVE NONLINEAR FILTERING

David W. Whitcombe

Aerospace Corporation

Prepared for:

Space and Missile Systems Organization

30 November 1972

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AIR FORCE REPORT NO.
SAMSO-TR-73-265

AEROSPACE REPORT NO.
TR-0073(3115)-2

Pseudo State Measurements Applied to Recursive Nonlinear Filtering

Prepared by D. W. WHITCOMBE
Guidance and Control Division

72 NOV 30

Engineering Science Operations
THE AEROSPACE CORPORATION

Prepared for SPACE AND MISSILE SYSTEMS ORGANIZATION
AIR FORCE SYSTEMS COMMAND
LOS ANGELES AIR FORCE STATION
Los Angeles, California

33

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) The Aerospace Corporation El Segundo, California	2a REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3 REPORT TITLE

PSEUDO STATE MEASUREMENTS APPLIED TO RECURSIVE NONLINEAR FILTERING

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

5 AUTHOR(S) (First name, middle initial, last name)

David W. Whitcombe

6 REPORT DATE 72 NOV 30	7a TOTAL NO. OF PAGES ~~30~~ 33	7b NO OF REFS 5

8a CONTRACT OR GRANT NO F04701-72-C-0073 b PROJECT NO. c d	9a ORIGINATOR'S REPORT NUMBER(S) TR-0073(3115)-2 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report) SAMSO-TR-73-265

10 DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Space and Missile Systems Organization Air Force Systems Command Los Angeles, California

13. ABSTRACT

"Pseudo state measurements" are constructed to make the measurement (geometry) model linear in the state. In the past, linear measurements have often proved to give better state estimates than nonlinear measurements. They are nonlinear functions of the actual measurement model bias parameters and are constructed to be linear functions of the state variables or to vanish in the absence of model or measurement error. Some examples of constructing pseudo state measurements are given in the paper. Recursive filter equations are derived using the pseudo state measurements and including colored (Markov) measurement noise and unestimated state and measurement model parameters. The filter estimates minimize the usual weighted least squares cost function with correlated state and pseudo state measurements. The filter is linear by construction. Higher order partial derivatives, if retained, would appear only in the computation of error variance and covariance matrices.

DD FORM 1473 (FACSIMILE)

UNCLASSIFIED
Security Classification

14

KEY WORDS

Estimation
Filtering
Kalman
Smoothing
Prediction
Recursive
Nonlinear

Distribution Statement (Continued)

Abstract (Continued)

FOREWORD

This report is published by The Aerospace Corporation, El Segundo, California, under Air Force Contract No. F04701-72-C-0073.

This report, which documents research carried out from July 1972 through January 1973, was submitted for review and approval on 30 May 1973 to Capt. Edward F. Stafford, SAMSO/DYGM.

Approved by

J. R. Allder

J. R. Allder, Director
Guidance and Flight Dynamics Subdivision
Guidance and Control Division
Engineering Science Operations

Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

Edward F. Stafford

EDWARD F. STAFFORD, Capt, USAF
Project Officer, Missile Guidance & Technology Div, Guidance Directorate

ABSTRACT

"Pseudo state measurements" are constructed to make the measurement (geometry) model linear in the state. In the past, linear measurements have often proved to give better state estimates than nonlinear measurements. They are nonlinear functions of the actual measurement model bias parameters and are constructed to be linear functions of the state variables or to vanish in the absence of model or measurement error. Some examples of constructing pseudo state measurements are given in the paper. Recursive filter equations are derived using the pseudo state measurements and including colored (Markov) measurement noise and unestimated state and measurement model parameters. The filter estimates minimize the usual weighted least squares cost function with correlated state and pseudo state measurements. The filter is linear by construction. Higher order partial derivatives, if retained, would appear only in the computation of error variance and covariance matrices.

CONTENTS

FIGURES

NOMENCLATURE

SCALARS

n	computation cycle at t_n
T	denotes transpose
t_n	time of n^{th} measurement
λ_n	weighted least squares cost function
$\overline{\lambda}_n$	$\lambda_n(\overline{X}_n)$
$< \cdot \cdot \cdot >$	expectation operator

VECTORS		DIMENSIONS
e_n	$X_n - \hat{X}_n$	n_1
$\ell_n(X_n)$	constructed linear functions of the state variables	n_5
N_n	error in Z_n	n_2
p	state unestimated parameters	n_3
$\overline{p}$	estimate of p	n_3
$\tilde{p}$	$p - \overline{p}$	n_3
T	unestimated measurement model parameters	n_4
$\overline{T}$	estimate of T	n_4
$\tilde{T}$	$T - \overline{T}$	n_4
U_n, W_n	white noise vectors	n_1, n_2
X_n	state vector at t_n	n_1
$\overline{X}_n$	state estimate at t_n	n_1
$\tilde{X}_n$	$X_n - \overline{X}_n$	n_1

Preceding page blank

$\hat{X}_n$	estimate of X_n using Z_n	n_1
Z_n	measurement at t_n	n_2
Z_n^*	$Z_n - N_n$	n_2
$\Delta\zeta_n$	$\zeta_n^* - \zeta_n$	n_5
ζ_n	$\zeta_n(Z_n, \overline{X}_n, \overline{T})$ pseudo state measurements	n_5
ζ_n^*	$\zeta_n(Z_n^*, X_n, T)$	n_5
$\varphi(X_{n-1}, p)$	state transition function	n_1

MATRICES

A	$[\Sigma - CM^{-1} C^T]^{-1}$	$n_5 \times n_5$
B	$\langle \tilde{T} \tilde{T}^T \rangle$	$n_4 \times n_4$
C_n	$-\langle \Delta\zeta_n \tilde{X}_n^T \rangle$	$n_5 \times n_1$
D	$[M - C^T \Sigma^{-1} C]^{-1}$	$n_1 \times n_1$
$\mathscr{D}$	$\partial\varphi/\partial p \vert \overline{p}, \hat{X}_{n-1}$	$n_1 \times n_3$
$\mathscr{E}_n$	$\partial\zeta_n/\partial\overline{T}$	$n_5 \times n_4$
F_n	$I + K_n \mathscr{H}_n - K_n \mathscr{L}_n$	$n_1 \times n_1$
G_n	$\langle e_n N_n^T \rangle$	$n_1 \times n_2$
$\mathscr{G}_n$	$\partial\zeta_n/\partial Z_n$	$n_5 \times n_2$
$\mathscr{H}_n$	$\partial\zeta_n/\partial\overline{X}_n$	$n_5 \times n_1$
I	identity matrix	. . .
J	$-[\Sigma - CM^{-1} C^T]^{-1} CM^{-1}$	$n_1 \times n_5$
$K_n, \mathscr{K}_n$	filter gain matrix	$n_1 \times n_5$
L_n	$\langle e_n \tilde{T}^T \rangle$	$n_1 \times n_4$
$\mathscr{L}_n$	$\partial \ell_n/\partial X_n$	$n_5 \times n_1$

M_n	$\langle \tilde{X}_n \tilde{X}_n^T \rangle$	$n_1 \times n_1$
P_n	$\langle e_n e_n^T \rangle$	$n_1 \times n_1$
Q_n	$\langle U_n U_n^T \rangle$	$n_1 \times n_1$
q	correlation of N_n with W_n	$n_2 \times n_2$
R_n	$\langle N_n N_n^T \rangle$	$n_2 \times n_2$
S	$\langle \tilde{p} \cdot \tilde{p}^T \rangle$	$n_3 \times n_3$
V_n	$\langle e_n \tilde{p}^T \rangle$	$n_1 \times n_3$
W_n	$\langle \tilde{X}_n N_n^T \rangle$	$n_1 \times n_2$
Y_n	$\langle \tilde{X}_n \tilde{T}^T \rangle$	$n_1 \times n_4$
ρ	correlation coefficients relating N_{n-1} to N_n	$n_2 \times n_2$
Σ_1	$\langle \Delta\zeta_n \Delta\zeta_n^T \rangle$	$n_5 \times n_5$
ϕn	$\partial\varphi/\partial X_n \mid \hat{X}_{n-1}, \overline{p}$	$n_1 \times n_1$

SECTION I

INTRODUCTION

Nonlinear problem geometries often assume that the measurements are related to the state using the nonlinear relation $Z_n = h_n(X_n) + N_n$. With this nonlinear form of the measurement model employed, the current optimal estimate can only approximately be represented as a linear combination of the measurements. The usual procedure is either to iterate the solution until acceptable accuracy is obtained or simply to ignore the partial derivatives of $h_n(X_n)$ that are higher than the first (Ref. 1). In this latter case, a direct computation of the errors caused by the omission of the second derivatives in some applications exceeds the measurement noise by an order of magnitude. Hence a serious convergence problem may arise with this formulation. Techniques for using the measurement model second derivatives are given in Ref. 2.

A new approach to nonlinear filtering with correlated measurement noise is presented in this paper. This approach, using pseudo state measurements, differs from the usual approach of the "extended Kalman" filter. The latter approach requires the computation of nonlinear residuals and uses nonlinear propagation of the state. The approach defined in this paper uses a nonlinear transformation of the actual measurements and the a priori state variables to obtain pseudo state measurements. The analysis in this paper is general in that all such functions are permitted, including functions that vanish in the noise-free case. The principal advantage results when the pseudo state measurements are constructed to be linear functions of the state variables.

All nonlinearity in the measurement model is restricted to the computation of the covariance and variance error matrices. These nonlinearities are unavoidable and do not appear to be a source of concern.

The usual weighted least squares cost function of correlated residuals is adopted as the cost function to be minimized. The use of pseudo state measurements has been tested to a limited extent by the author, using both simulated and actual data with good results.

The derivation using pseudo state measurements in this paper includes the derivation in Ref. 3 as a special case. The equations (see Appendix A) degenerate to the usual Kalman filter by setting S, o, $\mathscr{E}$, V, D, G, L, $\mathscr{H}$, and B to zero and $\mathscr{G}$ to I.

SECTION II

EXAMPLES OF PSEUDO STATE MEASUREMENTS

Pseudo state measurements ζ_n are constructed as nonlinear functions of the actual measurements Z_n and the a priori state estimate $\overline{X}_n$ as $\zeta_n = \zeta_n(Z_n, \overline{X}_n)$. It is convenient to construct pseudo state measurements to be linear functions of the state variables $\ell_n(X_n)$. The pseudo state residuals are then $\Delta r_n = \zeta_n(Z_n, \overline{X}_n) - \ell_n(\overline{X}_n)$, where $\ell_n(\overline{X}_n) = \mathscr{L}_n \overline{X}_n$ are linear functions of the state variables by construction.

Several examples will be given to illustrate the construction of pseudo state measurements. Consider that the state variables include X_1, X_2, X_3, $\dot{X}_1$, $\dot{X}_2$, and $\dot{X}_3, \ldots,$. It is possible that three simultaneous position measurements can be made such that X_1, X_2, and X_3 can be calculated from the three measurements. Examples of this include trilateration, where three simultaneous range measurements are made from three different stations. (In this context a station may denote either a ground station, a ship, or a satellite whose coordinates are known approximately.) Other examples can be constructed wherein exactly three position measurements involving range, angle, and/or direction cosines, etc., are measured from which X_1, X_2, and X_3 can be calculated in the ideal (error-free) case (Refs. 4 and 5).

A simple example occurs when range, azimuth, and elevation angles R, A, E are measured from a station. Then

$$\zeta_1 = R \cos A \cos E \qquad (\ell_1 = X_1)$$

$$\zeta_2 = R \sin A \cos E \qquad (\ell_2 = X_2)$$

$$\zeta_3 = R \sin E \qquad (\ell_3 = X_3)$$

are pseudo state measurements since these are nonlinear functions of the actual measurements. This is a special case since a priori state variables

were not required in the construction. In general, it is not possible to find a nonlinear transformation of the measurements that will be linear functions of the state variables. Consider the case where two angle measurements are made from a single station. Then the pseudo state measurements (PSMs) become

$$\zeta_1 = \overline{R} \cos A \cos E \qquad (\ell_1 = X_1)$$

$$\zeta_2 = \overline{R} \sin A \cos E \qquad (\ell_2 = X_2)$$

$$\zeta_3 = \overline{R} \sin E \qquad (\ell_3 = X_3)$$

where $\overline{R} = (\overline{X}_1^2 + \overline{X}_2^2 + \overline{X}_3^2)^{1/2}$ is constructed using the a priori state vector. If E, A are measured at several stations, then the same procedure may be followed for each station. This procedure extends the use of the a priori data to the measurement model. It may be preferable in this example to omit ζ_2 or ζ_3 since there are only two measurements. The use of only two PSMs may avoid some matrix inversion problems in the gain computation. An alternate construction of the PSM for angle only tracking data is as follows:

$$\zeta_1 = -\overline{X}_1 \sin A + \overline{X}_2 \cos A \qquad (\ell_1 = 0)$$

$$\begin{aligned} \zeta_2 = {} & \overline{X}_1 \sin E \cos A \\ & + \overline{X}_2 \sin E \sin A - \overline{X}_3 \cos E \qquad (\ell_2 = 0) \end{aligned}$$

This technique has the further advantage that $\partial\zeta/\partial\overline{X}_i$ are independent of $\overline{X}_i$.

In the case where R is measured at one or more stations, it is only necessary to construct $\overline{E}$, $\overline{A}$ from the a priori data in order to obtain ζ_1, ζ_2, and ζ_3 (or any one of the three). A similar procedure can be constructed for velocity, acceleration, and attitude measurements obtained in the sensor coordinate system. These measurements may be linearly related to the state vector.

For example, when the measurements are $\dot{R}$, $\dot{E}$, and $\dot{A}$, then the pseudo state velocity measurements may be constructed as

$$\zeta_i = (\dot{R},\ \dot{E},\ \dot{A};\ \overline{R},\ \overline{E},\ \overline{A}) \qquad (\ell_i = \dot{X}_i) \qquad (i = 1, 2, 3)$$

where $\overline{R}$, $\overline{E}$, and $\overline{A}$ are calculated from $\overline{X}$, $\overline{Y}$, and $\overline{Z}$. In a similar fashion, vehicle measurements of acceleration and attitude rate may be linearly related to the state vector.

SECTION III

FILTER COMPARISON

The pseudo state measurement filter (PSM) can be compared with the extended Kalman filter (EKF) in the following special case involving white measurement noise and n_ν unestimated parameters. The result is given below:

<u>EKF</u>

$$K_n = M_n \mathscr{L}_n^T (R_n + \mathscr{L}_n M_n \mathscr{L}_n^T)^{-1}$$

$$P_n = (I - K_n \mathscr{L}_n) M_n$$

$$\hat{X}_n = \bar{X}_n + K_n [\zeta_n - \ell_n(\bar{X}_n)]$$

<u>PSM</u>

$$K_n = M_n(\mathscr{L}_n - \mathscr{H}_n)^T [\mathscr{G}_n R_n \mathscr{G}_n^T + (\mathscr{L}_n - \mathscr{H}_n) M_n (\mathscr{L}_n - \mathscr{H}_n)^T]^{-1}$$

$$P_n = [I - K_n(\mathscr{L}_n - \mathscr{H}_n)] M_n$$

$$\hat{X}_n = \bar{X}_n + K_n(\zeta_n - \mathscr{L}_n \bar{X}_n)$$

where $M_n \quad \phi_n P_{n-1} \phi_n^T + Q_n$ and $\bar{X}_n$ are evaluated using the nonlinear formulation. The balance of the symbols are defined in Appendix A. Note that no errors are introduced in the PSM formulation as a result of neglecting second-order partial derivatives of $\ell_n(X_n)$. In one of the examples of PSMs, it follows that both $\mathscr{L}_n$ and $\mathscr{H}_n$ are constants. The EKF formulation suffers from the fact that $\mathscr{L}_n$ are not constants.

APPENDIX A

SUMMARY

The filtering concept may be explained with reference to the block diagram (Fig. A-1), which shows seven computational blocks denoted as:

Block 1. State propagation
Block 2. Pseudo measurements
Block 3. Set up computations
Block 4. Cost function
Block 5. Gain and covariance computation
Block 6. Step initialization
Block 7. Filter performance evaluation

The previous best estimate of the state vector is propagated in Block 1 to the time t_n when the measurements are made. Pseudo state measurements and pseudo state residuals are calculated in Block 2. The gain matrix K_n is computed recursively in Block 5. The gain matrix scales the pseudo state measurement residuals to obtain the improvement increment to be added to $\overline{X}_n$. The error covariance matrix of the estimate is also computed in Block 5.

The computation of the gain matrix requires that matrices M, R, Σ, and C be computed. This is performed in set up (Block 3).

Block 6, step initialization, is included so that the required input matrices to Block 5 (G, V, and L) may be computed recursively.

Blocks 4 and 7 are included so that the filter performance, including dynamics and measurement modeling, may be evaluated. Note that inclusion of these blocks is not required for the recursive filter operation. These may be included when the computational capability permits and when there is some question as to the adequacy of the modeling process.

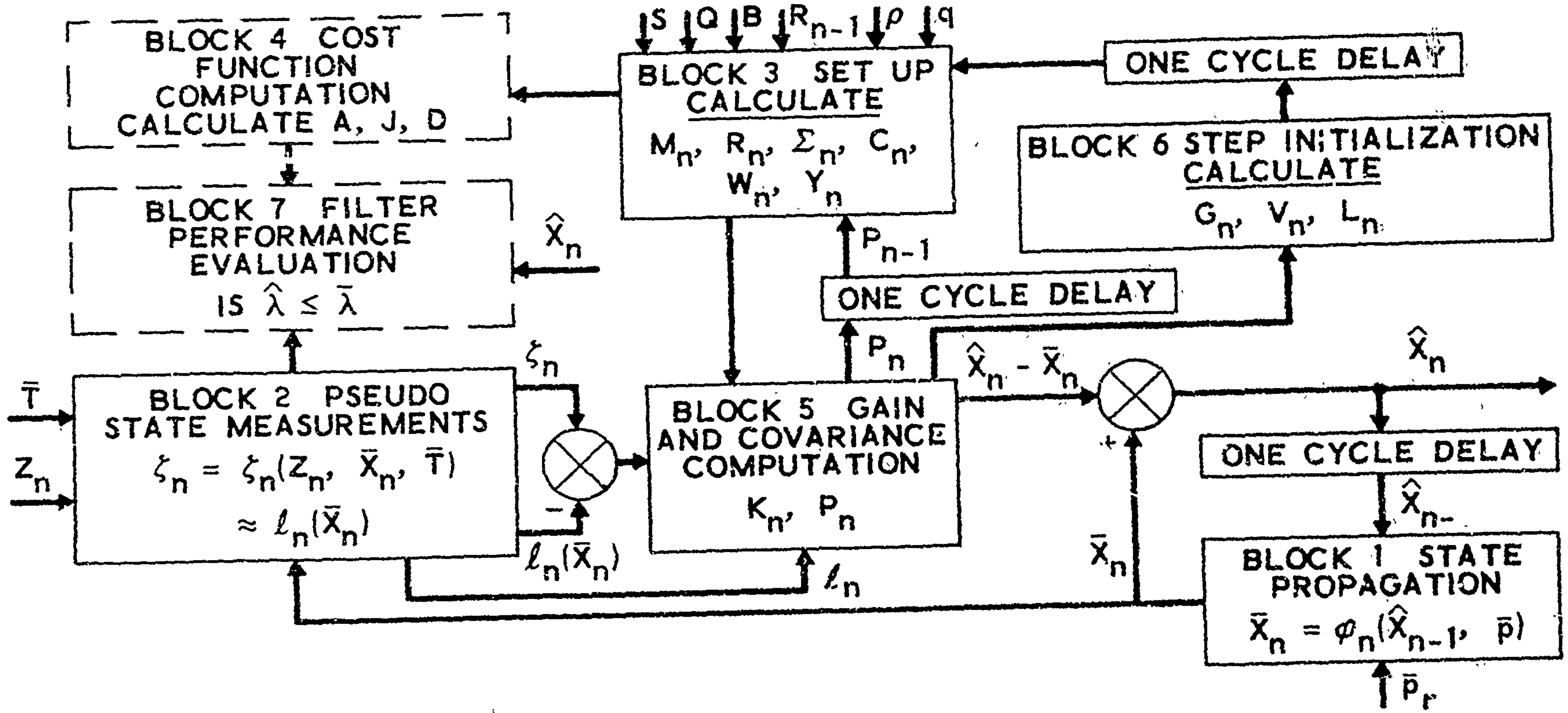

Figure A-1. System Block Diagram

Block 3 Equations

$$W_n = \phi_n G_{n-1} \rho^T$$

$$Y_n = \phi_n L_{n-1}$$

$$R_n = \rho R_{n-1} \rho^T + q q^T$$

$$M_n = \phi_n P_{n-1} \phi_n^T + \phi V_{n-1} \mathscr{D}^T + \mathscr{D} V_{n-1}^T \phi^T + \mathscr{D} S \mathscr{D}^T + Q_n$$

$$C_n = -\mathscr{H}_n M_n - \mathscr{G}_n W_n^T - \mathscr{E}_n Y_n^T$$

$$\Sigma_n = \mathscr{G}_n R_n \mathscr{G}_n^T + \mathscr{E}_n B \mathscr{E}_n^T + \mathscr{G}_n W_n^T \mathscr{H}_n^T + \mathscr{H}_n W_n \mathscr{G}_n^T + \mathscr{H}_n M_n \mathscr{H}_n^T + \mathscr{H}_n Y_n \mathscr{E}_n^T + \mathscr{E}_n Y_n^T \mathscr{H}_n^T$$

Block 5 Equations

$$K_n = (C_n^T + M_n \mathscr{L}_n^T) \cdot (\Sigma_n + \mathscr{L}_n C_n^T + C_n \mathscr{L}_n^T + \mathscr{L}_n M_n \mathscr{L}_n^T)^{-1}$$

$$\hat{X}_n = \overline{X}_n + K_n [\zeta_n - \ell_n(\overline{X}_n)]$$

$$P_n = (I - K_n \mathscr{L}_n) M_n - K_n C$$

Block 6 Equations

$$F_n = I + K_n \mathscr{H}_n - K_n \mathscr{L}_n$$

$$G_n = F_n W_n + K_n \mathscr{G}_n R_n$$

$$V_n = F_n \phi_n V_{n-1} + F_n \mathscr{D} S$$

$$L_n = F_n Y_n + K_n \mathscr{E}_n B$$

Block 7 Equations

$$\lambda = [\zeta_n - \ell_n(X_n)]^T A_n [\zeta_n - \ell_n(X_n)] + 2[\zeta_n - \ell_n(X_n)]^T J_n (X_n - \overline{X}) + (X_n - \overline{X}_n)^T D_n (X_n - \overline{X}_n)$$

APPENDIX B

DERIVATION

The filter equations will be derived with reference to Fig. A-1. The equations for each of the blocks will be derived.

B. 1 Block 1 - State Propagation

The dynamics of the state vector propagation are given by the non-linear vector function

$$X_n = \varphi(X_{n-1}, p) + U_n \qquad (1-1)$$

It is assumed that the best estimates $\hat{X}_{n-1}$ and $\overline{p}$ are available such that

$$\overline{X}_n = \varphi(\hat{X}_{n-1}, \overline{p}_r) \qquad (1-2)$$

B. 2 Block 2 - Pseudo State Measurements

Let Z_n denote the actual measurements = $Z_n^* - N_n$. These measurements can be correlated as a first-order Markov process as

$$N_n = \rho N_{n-1} + q W_n \qquad (2-1)$$

The development in this report can be extended to the n^{th}-order Markov process. Pseudo state measurements are now constructed as nonlinear functions of the actual measurements and the a priori state estimate $\overline{X}_n$. The pseudo state measurements then become linear functions of the state variables; i. e.

$$\zeta_n = \zeta_n(Z_n, \overline{X}_n, \overline{T}) \qquad (2-2)$$

denotes the pseudo state measurements. These are constructed to be exactly linear functions of the state variables except for the errors $\Delta\zeta_n$ that result from the measurement noise N_n, the state error estimate $\tilde{x}_n$, and the measurement model errors $\tilde{T}$. Hence, in the equation

$$\zeta_n = \ell_n(X_n) - \Delta\zeta_n \qquad (2\text{-}3)$$

ℓ_n denotes a linear function of the state variables. In practice, an attempt is made for ℓ_n to be identical with one or more of the state vector components. Pseudo state measurement residuals are then obtained (Fig. A-1) as $\zeta_n - \ell_n(\overline{X}_n)$.

B.3 Block 3 - Set Up

The computation of the cost function requires that M_n, R_n, Σ_n, and C_n be calculated. This is accomplished by using the inputs to Block 3 shown in Fig. A-1.

The computation of R_n, using Eq. (2-1), is obtained as

$$\begin{aligned} R_n &= \langle N_n N_n^T \rangle \\ &= \langle (\rho N_{n-1} + q W_n)(N_{n-1}^T \rho^T + W_n^T q^T) \rangle \\ &= \rho \langle N_{n-1} N_{n-1}^T \rangle \rho^T + q \langle W_n W_n^T \rangle q^T \\ &= \rho R_{n-1} \rho^T + q q^T \end{aligned} \qquad (3\text{-}1)$$

where $\langle W_n W_n^T \rangle = I$.

The computation of M_n requires the evaluation of $\tilde{X}_n$, which results from Eqs. (1-1) and (1-2) as

$$
\begin{aligned}
\widetilde{X}_n &= X_n - \overline{X}_n \\
&= \varphi(X_{n-1}, p) + U_n - \varphi(\hat{X}_{n-1}, \overline{p}) \\
&= \varphi(\hat{X}_{n-1} + e_{n-1}, \overline{p} + \widetilde{p}) + U_n - \varphi(\hat{X}_{n-1}, \overline{p}) \\
&= \phi_n\, e_{n-1} + \mathscr{D}_n\, \widetilde{p} + U_n
\end{aligned} \tag{3-2}
$$

where the expansion terms higher than the first degree were ignored.

The computation of Σ_n and C_n require the evaluation of $\Delta\zeta_n$, as defined in Eq. (2-3). Using Eq. (2-2), it follows that

$$
\begin{aligned}
\Delta\zeta_n &= \zeta_n(Z_n^*, X_n, T) - \zeta_n(Z_n, \overline{X}_n, \overline{T}) \\
&= \zeta_n(Z_n + N_n, \overline{X}_n + \widetilde{X}_n, \overline{T} + \widetilde{T}) - \zeta_n(Z_n, \overline{X}_n, \overline{T})
\end{aligned} \tag{3-3}
$$

The result of expanding the first term in a Taylor's series and retaining only the linear terms is

$$
\Delta\zeta_n = \mathscr{G}_n N_n + \mathscr{H}_n \widetilde{X}_n + \mathscr{E}_n \widetilde{T} \tag{3-4}
$$

where the partial derivatives are defined as

$$
\mathscr{G}_n = \left.\frac{\partial \zeta_n}{\partial Z_n}\right|_{Z_n, \overline{X}_n, \overline{T}}
$$

$$
\mathscr{H}_n = \left.\frac{\partial \zeta_n}{\partial X_n}\right|_{Z_n, \overline{X}_n, \overline{T}}
$$

$$
\mathscr{E}_n = \left.\frac{\partial \zeta_n}{\partial T}\right|_{Z_n, \overline{X}_n, \overline{T}}
$$

In the computation of M_n, Eq. (3-2) is used as

$$
\begin{aligned}
M_n &= \langle \widetilde{X}_n \widetilde{X}_n^T \rangle \\
&= \langle (\phi_n e_{n-1} + \mathscr{D}\widetilde{p} + U_n)(e_{n-1}^T \phi_n^T + \widetilde{p}^T \mathscr{D}^T + U_n^T) \rangle \\
&= \phi_n P_{n-1} \phi_n^T + \phi_n V_{n-1} \mathscr{D}^T + \mathscr{D} V_{n-1}^T \phi^T \\
&\quad + \mathscr{D} S \mathscr{D}^T + Q_n
\end{aligned}
\tag{3-5}
$$

In the computation of C_n, Eqs. (3-2) and (3-4) are used as

$$
\begin{aligned}
C_n &\equiv -\langle \Delta \zeta_n \widetilde{X}_n^T \rangle \\
&= -\mathscr{G}_n \langle N_n \widetilde{X}_n^T \rangle - \mathscr{H}_n \langle \widetilde{X}_n \widetilde{X}_n^T \rangle - \mathscr{E}_n \langle \widetilde{T} X_n^T \rangle
\end{aligned}
$$

but from Eqs. (3-2) and (2-1), it follows that

$$
\begin{aligned}
W_n \equiv \langle \widetilde{X}_n N_n^T \rangle &= \langle (\phi_n e_{n-1} + \mathscr{D}\widetilde{p} + U_n)(N_{n-1}^T \rho^T + W_n^T q T) \rangle \\
&= \phi_n \langle e_{n-1} N_{n-1}^T \rangle \rho^T = \phi_n G_{n-1} \rho^T
\end{aligned}
\tag{3-6}
$$

$$
Y_n \equiv \langle \widetilde{X}_n \widetilde{T}^T \rangle = \phi_n L_{n-1} \tag{3-7}
$$

Hence

$$
C_n = -\mathscr{H}_n M_n - \mathscr{G}_n W_n^T - \mathscr{E}_n Y_n^T \tag{3-8}
$$

where it is assumed that

$$
0 = \langle \widetilde{p} N_n^T \rangle = \langle \widetilde{p} \widetilde{T}^T \rangle = \langle \widetilde{T} N_n^T \rangle
$$

The computation of Σ_n is obtained by using Eqs. (3-4), (3-2), and (2-1) as

$$
\begin{aligned}
\Sigma_n &= \langle \Delta\zeta_n \, \Delta\zeta_n^T \rangle \\
&= \langle (\mathscr{G}_n N_n + \mathscr{H}_n \tilde{X}_n + \mathscr{E}_n \tilde{T}) (N_n^T \mathscr{G}_n^T + \tilde{X}_n^T \mathscr{H}_n^T + \tilde{T}^T \mathscr{E}_n^T) \rangle \\
&= \mathscr{G}_n R_n \mathscr{G}_n^T + \mathscr{G}_n W_n^T \mathscr{H}_n^T + \mathscr{H}_n W_n \mathscr{G}_n^T + \mathscr{H}_n M_n \mathscr{H}_n^T \\
&\quad + \mathscr{H}_n Y_n \mathscr{E}_n^T + \mathscr{E}_n Y_n^T \mathscr{H}_n^T + \mathscr{E}_n B \mathscr{E}_n^T
\end{aligned}
\tag{3-9}
$$

B.4 Block 4 - The Cost Function

The cost function is constructed to minimize the correlated weighted least squares of the sum vector, consisting of the pseudo state measurement residuals and the a priori state residuals, as

$$
\begin{pmatrix} \zeta_n - \ell_n(X_n) \\ X_n - \bar{X}_n \end{pmatrix}
$$

i. e., the two vectors are joined to form a single vector consisting of the ordered sum of all components. The cost function then becomes

$$
\begin{aligned}
\lambda_n &= \left(\zeta_n - \ell_n(X_n)\right)^T A_n \left(\zeta_n - \ell_n(X_n)\right) + 2\left(\zeta_n - \ell_n(X_n)\right)^T J_n(X_n - \bar{X}_n) \\
&\quad + (X_n - \bar{X}_n)^T D_n (X_n - \bar{X}_n)
\end{aligned}
\tag{4-1}
$$

where the quantities A, J, and D are defined as

$$
\begin{pmatrix} A_n & J_n \\ J_n^T & D_n \end{pmatrix} = \begin{pmatrix} \Sigma_n & C_n \\ C_n^T & M_n \end{pmatrix}^{-1}
\tag{4-2}
$$

and

$$\Sigma_n \equiv \left\langle \left(\zeta_n - \ell_n(X_n)\right) \left(\zeta_n - \ell_n(X_n)\right)^T \right\rangle = \langle \Delta\zeta_n \, \Delta\zeta_n^T \rangle$$

$$C_n \equiv \left\langle \left(\zeta_n - \ell_n(X_n)\right) (X_n - \overline{X}_n) \right\rangle = -\langle \Delta\zeta_n \, \tilde{X}_n^T \rangle$$

$$M_n \equiv \langle (X_n - \overline{X}_n)(X_n - \overline{X}_n)^T \rangle = \langle \tilde{X}_n \, \tilde{X}_n^T \rangle$$

are each calculated in Block 3. Note that if $\Delta\zeta_n$ and $X_n - \overline{X}_n$ are Gaussian, the cost function becomes the conditional mean and the estimate is considered optimal in the usually accepted sense. However, these residuals need not be Gaussian for the purposes of this report. The cost function is then simply the weighted least squares cost function. Note also that this formulation of the cost function is such that, given Gaussian variables, the distribution function of the cost function is chi-square with $n_1 + n_2$ degrees of freedom and that these data could be transformed to orthogonal variables. If this were done, each item in the sum of squares would be weighted inversely as its variance.

Expressions for A_n, D_n, and J_n will be calculated. Multiplying both sides of Eq. (4-2) together and equating to the identity matrix results in the following equations

$$\begin{aligned} \Sigma A + CJ^T &= I \\ \Sigma J + CD &= O \\ C^T A + MD &= I \\ C^T A + MJ^T &= O \end{aligned} \qquad (4\text{-}3)$$

where the n subscripts have been omitted for ease in typing. The result of solving for A, D, and J is

$$A = [\Sigma - CM^{-1} C^T]^{-1} = \Sigma^{-1} + \Sigma^{-1} CDC^T \Sigma^{-1} \qquad (4\text{-}4)$$

$$D = [M - C^T \Sigma^{-1} C]^{-1} = M^{-1} + M^{-1} C^T ACM^{-1} \tag{4-5}$$

and

$$\begin{aligned} J &= -[\Sigma - C M^{-1} C^T]^{-1} C M^{-1} \\ &= -\Sigma^{-1} C[M - C^T \Sigma^{-1} C]^{-1} \end{aligned} \tag{4-6}$$

There are two solution forms for each expression because of the redundancy of Eqs. (4-3) (4 equations with only 3 unknowns). The user should choose the combination that provides the simplest computer program.

B.5 Block 5 - The Gain Computation

The optimal estimate is the one that minimizes the cost function obtained in Eq. (4-1). This estimate results from taking the gradient of Eq. (4-1)

$$\begin{aligned} \frac{\partial \lambda_n}{\partial X_n} &= -2\mathscr{L}_n^T A_n\left(\zeta_n - \ell_n(X_n)\right) + 2 J_n^T \left(\zeta_n - \ell_n(X_n)\right) \\ &\quad - 2\mathscr{L}_n^T J_n(X_n - \overline{X}_n) + 2 D_n(X_n - \overline{X}_n) \\ &= 0 \text{ when } X_n = \hat{X}_n \end{aligned} \tag{5-1}$$

The result of solving Eq. (5-1) for $\hat{X}_n$ is

$$(D_n - \mathscr{L}_n^T J_n)(\hat{X}_n - \overline{X}_n) = (\mathscr{L}_n^T A_n - J_n^T)\left(\zeta_n - \ell_n(\hat{X}_n)\right) \tag{5-2}$$

Now, since $\ell_n(\hat{X}_n)$ is a linear function, it may be written as

$$\ell_n(\hat{X}_n) \equiv \ell_n(\overline{X}_n) + \mathscr{L}_n(\hat{X}_n - \overline{X}_n) \tag{5-3}$$

and then Eq. (5-2) may be written as

$$\hat{X}_n = \bar{\bar{X}}_n + K_n[\zeta_n - \ell_n(\bar{X}_n)] \tag{5-4}$$

where

$$K_n = [D_n + \mathscr{L}_n^T A_n \mathscr{L}_n - J_n^T \mathscr{L}_n - \mathscr{L}_n^T J_n]^{-1} [\mathscr{L}_n^T A_n - J_n^T] \tag{5-5}$$

Note that the gain matrix K_n consists only of constants, i.e., K_n is independent of the state and measurement residuals.

The gain computed in Eq. (5-5) gives an estimate $\hat{X}_n$ that minimizes the weighted least squares cost function of Eq. (4-1). It will now be shown that the same gain K_n also minimizes each of the diagonal elements of the error covariance matrix P_n. This assumes that the minimum variance estimate is a linear function of the pseudo state measurement residuals; i.e., assume that

$$\hat{X}_n = \bar{X}_n + \mathscr{K}_n[\zeta_n - \ell_n(\bar{X}_n)] \tag{5-6}$$

Then the estimation error is obtained by adding and subtracting X_n from both sides of Eq. (5-6) to obtain

$$\begin{aligned} e_n = \tilde{X}_n - \mathscr{K}_n[\, & \zeta_n(Z_n, \bar{X}_n, \bar{T}) - \zeta_n(Z_n^*, X_n, T) \\ & + \mathscr{L}_n(X_n) - \mathscr{L}_n(\bar{X}_n)] \end{aligned} \tag{5-7}$$

Note that Eq. (5-7) agrees with Eq. (5-4) since the identity

$$\mathscr{L}_n(X_n) - \zeta_n(Z_n^*, X_n, T) \equiv 0$$

has been added within the brackets. Hence, Eq. (5-7) may be written using Eq. (3-4) as

$$e_n = \tilde{X}_n + \mathscr{K}_n[\Delta\zeta_n - \mathscr{L}_n \tilde{X}_n] \tag{5-8}$$

The error covariance matrix of the estimate is $P_n \equiv \langle e_n e_n^T \rangle$. It is easily shown by direct differentiation that each of the diagonal terms in P_n is minimized by choosing $\mathscr{K}_n$ to satisfy the Weiner Hopf equation as

$$\langle \widetilde{X}_n (\Delta \zeta_n - \mathscr{L}_n \widetilde{X}_n)^T \rangle = -\mathscr{K}_n \langle (\Delta \zeta_n - \mathscr{L}_n \widetilde{X}_n)(\Delta \zeta_n - \mathscr{L}_n \widetilde{X}_n)^T \rangle \tag{5-9}$$

An interesting derivation of the Wiener Hopf equation is given by Johnson (Ref. 3). Equation (5-9) may be expanded and the definitions of the covariance functions used to obtain

$$C_n^T + M_n \mathscr{L}_n^T = \mathscr{K}_n [\Sigma_n + \mathscr{L}_n C_n^T + C_n \mathscr{L}_n^T + \mathscr{L}_n M_n \mathscr{L}_n^T] \tag{5-10}$$

It will now be shown that the $\mathscr{K}_n$ defined in Eq. (5-10) is exactly the same as the K_n specified in Eq. (5-5); i.e.,

$$K_n \equiv \mathscr{K}_n$$

This is done by substituting K_n from Eq. (5-5) for $\mathscr{K}_n$ in Eq. (5-10). The result of this substitution is

$$\begin{aligned} &[D + \mathscr{L}^T A \mathscr{L} - J^T \mathscr{L} - \mathscr{L}^T J] [M \mathscr{L}^T + C^T] \\ &\overset{?}{=} [\mathscr{L}^T A - J^T] [\Sigma + C \mathscr{L}^T + \mathscr{L} C^T + \mathscr{L} M \mathscr{L}^T] \end{aligned} \tag{5-11}$$

It is easily shown, using Eqs. (4-3), (4-4), (4-5), and (4-6), that Eq. (5-11) is an identity. The proof is simplified by noting that corresponding terms involving similar groupings of $\mathscr{L}$ and $\mathscr{L}^T$ must vanish identically. Hence it has been shown that the weighted least squares estimate is identical with the linear minimum variance estimate and that the simpler expression may be used for the gain computation as

$$K_n = [C_n^T + M_n \mathscr{L}_n^T] \cdot [\Sigma_n + \mathscr{L}_n C_n^T + C_n \mathscr{L}_n^T + \mathscr{L} M_n \mathscr{L}^T]^{-1} \tag{5-12}$$

In the computation of P_n, Eq. (5-8) is used as

$$e_n = \tilde{X}_n + K_n \Delta r_n \tag{5-13}$$

where

$$\Delta r_n \equiv \Delta \zeta_n - \mathscr{L}_n \tilde{X}_n \tag{5-14}$$

Then

$$\begin{aligned} P_n &\equiv \langle (\tilde{X}_n + K_n \Delta r_n) (\tilde{X}_n + K_n \Delta r_n)^T \rangle \\ &= \langle \tilde{X}_n \tilde{X}_n^T \rangle + K_n \langle \Delta r_n X_n^T \rangle + \langle X_n \Delta r_n^T \rangle K_n^T + K_n \langle \Delta r \Delta r^T \rangle K_n^T \end{aligned} \tag{5-15}$$

where

$$\begin{aligned} \langle \Delta r_n \tilde{X}_n \rangle &= \langle (\Delta \zeta_n - \mathscr{L}_n \tilde{X}_n) X_n^T \rangle \\ &= \langle \Delta \zeta_n \tilde{X}_n^T \rangle - \mathscr{L}_n \langle \tilde{X}_n \tilde{X}_n^T \rangle \\ &= -C_n - \mathscr{L}_n M_n \end{aligned} \tag{5-16}$$

and

$$\begin{aligned} \langle \Delta r_n \Delta r_n^T \rangle &= \langle (\Delta \zeta_n - \mathscr{L}_n \tilde{X}_n) (\Delta \zeta_n - \mathscr{L}_n \tilde{X}_n)^T \rangle \\ &= \Sigma_n + \mathscr{L}_n C_n^T + C_n \mathscr{L}_n^T + \mathscr{L}_n M_n \mathscr{L}^T \end{aligned} \tag{5-17}$$

Equation (5-15) may be simplified by noting that the Wiener Hopf equation (5-9) requires

$$\langle \tilde{X}_n \Delta r_n^T \rangle = -K_n \langle \Delta r_n \Delta r_n^T \rangle \tag{5-18}$$

Hence, P_n simplifies to

$$\begin{aligned} P_n &= \langle \tilde{X}_n \tilde{X}_n^T \rangle + K_n \langle \Delta r_n \tilde{X}_n^T \rangle \\ &= M_n - K_n [C_n + \mathscr{L}_n M_n] \qquad (5\text{-}19) \\ &= [I - K_n \mathscr{L}_n] M_n - K_n C \end{aligned}$$

P_n may also be computed using Eq. (5-15), which does not require that the filter satisfy the Wiener Hopf equation. The result is

$$\begin{aligned} P_n = M_n - K_n [C_n + \mathscr{L}_n M_n] - [C_n^T + M_n \mathscr{L}^T] K_n^T \\ + K_n [\Sigma_n + \mathscr{L}_n C_n^T + C_n \mathscr{L}_n^T + \mathscr{L}_n M_n \mathscr{L}_n^T] K_n^T \end{aligned} \qquad (5\text{-}20)$$

B.6 Block 6 - Step Initialization

The quantities G_n, V_n, and L_n are computed in Block 6. These quantities are required for the set up computations in Block 3.

The computation of these three covariance matrices requires computation of the estimation error. This is obtained in Eq. (5-12) as

$$e_n = \tilde{X}_n + K_n [\Delta \zeta_n - \mathscr{L}_n \tilde{X}_n] \qquad (6\text{-}1)$$

The use of Eq. (3-3) permits Eq. (6-1) to be written as

$$\begin{aligned} e_n &= \tilde{X}_n + K_n [\mathscr{G}_n N_n + \mathscr{H}_n \tilde{X}_n + \mathscr{E}_n \tilde{T} - \mathscr{L}_n \tilde{X}_n] \\ &= [I + K_n \mathscr{H}_n - K_n \mathscr{L}_n] \tilde{X}_n + K_n \mathscr{G}_n N_n + K_n \mathscr{E}_n \tilde{T} \end{aligned} \qquad (6\text{-}2)$$

$$= F_n \tilde{X}_n + K_n \mathscr{G}_n N_n + K_n \mathscr{E}_n \tilde{T} \qquad (6\text{-}3)$$

where

$$F_n \equiv I + K_n \mathscr{H}_n - K_n \mathscr{L}_n \tag{6-4}$$

Then G_n is obtained by using its definition and Eq. (3-5) as

$$G_n \equiv \langle e_n N_n^T \rangle$$

$$= \langle (F_n \tilde{X}_n + K_n \mathscr{G}_n N_n + K_n \mathscr{E}_n \tilde{T}) \cdot N_n^T \rangle \tag{6-5}$$

$$= F \langle \tilde{X}^n N_n^T \rangle + K_n \mathscr{G}_n \langle N_n N_n^T \rangle$$

$$= F W_n + K_n \mathscr{G}_n R_n \tag{6-6}$$

The matrix V_n is obtained using Eqs. (6-3) and (3-2) as

$$V_n \equiv \langle e_n \tilde{p}^T \rangle$$

$$= \langle (F_n \tilde{X}_n + K_n \mathscr{G}_n N_n + K_n \mathscr{E}_n \tilde{T}_x) \tilde{p}^T \rangle$$

$$= F_n \langle (\phi_n e_{n-1} + \mathscr{D} \tilde{p} + U_n) \tilde{p}^T \rangle \tag{6-7}$$

$$= F_n \phi_n \langle e_{n-1} \tilde{p}^T \rangle + F_n \mathscr{D} \langle \tilde{p} \tilde{p}^T \rangle$$

$$= F_n \phi_n V_{n-1} + F_n \mathscr{D} S$$

The matrix L_n is computed using Eqs. (6-3) and (3-7) as

$$\begin{aligned} L_n &\equiv \langle e_n \tilde{T}^T \rangle \\ &= \langle (F_n \tilde{X}_n + K_n \mathscr{G}_n N_n + K_n \mathscr{E}_n \tilde{T}) \tilde{T}^T \rangle \\ &= F_n \langle \tilde{X}_n \tilde{T}^T \rangle + K_n \mathscr{E}_n \langle \tilde{T} \tilde{T}^T \rangle \\ &= F_n Y_n + K_n \mathscr{E}_n B \end{aligned} \tag{6-8}$$

B.7 Block 7 - Filter Performance

The filter performance may be verified by computing the cost function given as Eq. (4-1) evaluated at $X_n = \overline{X}_n$ and also at $\hat{X}_n$. The equations are

$$\overline{\lambda}_n = \left(\zeta_n - \ell_n(\overline{X}_n)\right)^T A_n \left(\zeta_n - \ell_n(\overline{X}_n)\right) \tag{7-1}$$

and

$$\begin{aligned} \hat{\lambda}_n = &\left(\zeta_n - \ell_n(\hat{X}_n)\right)^T A_n \left(\zeta_n - \ell_n(\hat{X}_n)\right) + 2\left(\zeta_n - \ell_n(\hat{X}_n)\right)^T J_n(\hat{X}_n - \overline{X}_n) \\ &+ (\hat{X}_n - \overline{X}_n)^T D_n(\hat{X}_n - \overline{X}_n) \end{aligned} \tag{7-2}$$

It is necessary to verify that

$$\hat{\lambda}_n \leq \overline{\lambda}_n \tag{7-3}$$

in order to justify processing the measurements. In the event that this condition is not satisfied, the basic real-world dynamics and measurement models should be checked. The computer program should come to a halt whenever Eq. (7-3) is not satisifed.

REFERENCES

1. J. L. LeMay, "Theory and Applications of Linear Estimation Theory," Instrument Society of America, Preprint No. 3.2-1-65, Oct. 1965.

2. R. S. Bucy and P. D. Joseph, Filtering for Stochastic Processes with Applications to Guidance, Interscience Publishers, New York, 1968, Chapter 8.

3. D. J. Johnson, A General Linear Sequential Filter," AIAA Journal, Vol. 8, No. 9, Sept. 1970, pp. 1605-1608.

4. J. A. Henz, A Modified Kalman Filter for Radio Guidance Application, Report No. TR-0158(3124)-1, Aerospace Corp., El Segundo, Calif., Dec. 1967.

5. R. K. Mehra, "A Comparison of Two Nonlinear Filters for Ballistic Trajectory Estimation," Proceedings of the Symposium on Nonlinear Estimation Theory and Its Applications, IEEE No. 70C 66-AC, Sept. 1970.